

Spatial and Cross-Modal Attention Alter Responses to Unattended Sensory Information in Early Visual and Auditory Human Cortex

Vivian M. Ciaramitaro,^{1,2} Giedrius T. Buračas,³ and Geoffrey M. Boynton^{1,3}

¹Systems Neurobiology Laboratory, The Salk Institute for Biological Studies; and ²Department of Psychology and ³Center for Functional Magnetic Resonance Imaging, University of California San Diego, La Jolla, California

Submitted 22 May 2007; accepted in final form 17 August 2007

Ciaramitaro VM, Buračas GT, Boynton GM. Spatial and cross-modal attention alter responses to unattended sensory information in early visual and auditory human cortex. *J Neurophysiol* 98: 2399–2413, 2007. First published August 22, 2007; doi:10.1152/jn.00580.2007. Attending to a visual or auditory stimulus often requires irrelevant information to be filtered out, both within the modality attended and in other modalities. For example, attentively listening to a phone conversation can diminish our ability to detect visual events. We used functional magnetic resonance imaging (fMRI) to examine brain responses to visual and auditory stimuli while subjects attended visual or auditory information. Although early cortical areas are traditionally considered unimodal, we found that brain responses to the same ignored information depended on the modality attended. In early visual area V1, responses to ignored visual stimuli were weaker when attending to another visual stimulus, compared with attending to an auditory stimulus. The opposite was true in more central visual area MT+, where responses to ignored visual stimuli were weaker when attending to an auditory stimulus. Furthermore, fMRI responses to the same ignored visual information depended on the location of the auditory stimulus, with stronger responses when the attended auditory stimulus shared the same side of space as the ignored visual stimulus. In early auditory cortex, responses to ignored auditory stimuli were weaker when attending a visual stimulus. A simple parameterization of our data can describe the effects of redirecting attention across space within the same modality (spatial attention) or across modalities (cross-modal attention), and the influence of spatial attention across modalities (cross-modal spatial attention). Our results suggest that the representation of unattended information depends on whether attention is directed to another stimulus in the same modality or the same region of space.

INTRODUCTION

Each of our sensory modalities affords us a unique experience of the world, encoding sensory information along different dimensions for a given object or event. The information available across different modalities often correlates in a meaningful way and is integrated into a unified cross-modal percept. However, when cross-modal information is not correlated it may be necessary to filter out information in one modality while attending to another modality. For example, when talking on the telephone, salient changes in the visual scene can be rendered unnoticeable (Strayer and Johnston 2001). Similarly, it is common to be unaware of auditory information when engaged in a visual task, such as reading. These effects are subjectively powerful and brain imaging studies have started to consider how attention influences the representation not only of

attended and task-relevant information, but also of unattended, task-irrelevant information in a different modality (e.g., Weissman et al. 2004).

What has become increasingly clear is that cross-modal influences are observed not only in higher-level multimodal areas where early sensory areas converge. Rather, a growing body of evidence has highlighted the presence of cross-modal influences in early sensory areas (for reviews see Driver and Spence 1998, 2000; Eimer and Driver 2001; Ghazanfar and Schroeder 2006; Macaluso and Driver 2005; Shams and Shimojo 2001).

Within a given sensory modality, such as the visual modality, attention has been shown to modulate neuronal responses in early visual areas to visual stimuli based on what regions of space (*spatial attention*) and what visual features (*feature-based attention*) subjects are asked to attend (reviewed in Treue 2001). Studies have examined the distinct contributions of these different types of attentional selection on visual psychophysical performance (e.g., Baldassi and Vergheze 2005; Liu et al. 2007; Shih and Sperling 1996; Theeuwes 1989) and on responses in early visual areas, such as V4 or MT (Beauchamp et al. 1997; Hayden and Gallant 2005).

Similarly, in the auditory modality, responses in early auditory areas have been shown to be modulated by spatial attention (e.g., Alho et al. 1999; Jancke et al. 2003). To our knowledge, no study to date has yet considered differences in auditory cortex modulation as a function of feature-based attention, unless we think of space as a feature, in which case no differential modulation has been found based on the attended features (Zatorre et al. 1999).

Neuronal responses to sensory stimuli can be influenced not only by the spatial location or feature attended, but also by the modality of the attended information (*cross-modal attention*; reviewed in Driver and Spence 2000; Macaluso and Driver 2005; see also Shomstein and Yantis 2004). What are the distinct contributions of cross-modal versus spatial attention? It is known that cross-modal attention effects can be modulated by spatial factors. Within early sensory areas, redirecting spatial attention in one modality has been shown to influence responses to unattended stimuli in another modality (reviewed in Macaluso and Driver 2005), even when the stimuli across modalities are spatially misaligned (Murray et al. 2005) and even at a fine spatial scale within a hemifield (Eimer et al. 2004).

Address for reprint requests and other correspondence: V. M. Ciaramitaro, The Salk Institute for Biological Studies, Systems Neurobiology Laboratory-B, 10010 North Torrey Pines Rd., La Jolla, CA 92037-1099 (E-mail: vivian@salk.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Event-related potential (ERP) evidence from Eimer and colleagues (2004) shows that redirecting spatial attention to auditory stimuli enhances the N1 component of visual responses to ignored visual stimuli presented at the same spatial location. Given that the N1 component is thought to be generated in lateral occipitotemporal areas, this suggests that such spatial influences on cross-modal attention are present early in visual processing. However, it is unclear whether such spatial effects would be found as early as V1. Furthermore, given that spatial attention effects within the visual modality are present but weaker in V1, would cross-modal spatial effects also be weaker in V1?

Our goal was to determine the relative contributions of cross-modal and spatial attention for stimuli that were always ignored and not optimized for cross-modal integration, under conditions where all stimuli were presented concurrently, maximizing competitive interactions across modalities and across space. We compared the effects of spatial, cross-modal, and cross-modal spatial attention within early visual and auditory cortex. In particular, we examined how the same unattended task-irrelevant information would be represented in the brain when redirecting attention across space or across modalities and how such effects vary across early sensory areas. Functional magnetic resonance imaging (fMRI) responses were measured to stimuli that were: 1) attended; 2) ignored, with attention directed to a stimulus in the same modality, ipsilateral to the cortical area of interest; 3) ignored, with attention directed to a contralateral stimulus in a different modality; and 4) ignored, with attention directed to an ipsilateral stimulus in a different modality.

We hypothesize that spatial attention and cross-modal attention act independently to modulate neuronal responses. A simple parameterization of our data predicts our results and estimates the individual contributions of spatial attention and cross-modal attention. We found that whereas spatial attention factors are stronger than cross-modal factors within V1, the relative influence of cross-modal attention factors increases along the visual hierarchy. In auditory cortex, cross-modal attention factors also increase from primary to secondary auditory cortex, with a trend for cross-modal factors to be stronger than spatial factors. Furthermore, we found that cross-modal attention factors depended on whether the ignored and attended stimuli shared a common feature, such as being lateralized to the same side of space, and that such cross-modal spatial factors did not vary across either early visual or early auditory areas.

METHODS

Subject selection

Six human subjects, three females and three males (ages 23–32 yr), participated in both psychophysical and fMRI experiments. Subjects were undergraduate, graduate, or postdoctoral students recruited from the Salk Institute or the University of California San Diego community. Experiments were reviewed and approved by the Salk Institute Human Subjects Review Board. Written informed consent was obtained from all subjects. Subjects had normal or corrected-to-normal vision and could detect the range of auditory frequencies used in our experiment when presented monaurally. Two of our six subjects were left-handed. No subject had a psychiatric or neurological history. Although our study included only six subjects, each subject was well

trained on the task before scanning and each subject completed 48 scans (29,952 behavioral trials).

Stimulus presentation

Visual and auditory stimuli were generated with an Apple Macintosh computer (Powerbook G3 processor, 300 MHz) using the Psychophysics Toolbox (Brainard 1997; Pelli 1997), Video Toolbox, and the Matlab (v5.2) programming language. Viewing conditions were duplicated between the scanner and the testing room in the laboratory. The main difference was the subject's body position and the absence of auditory scanner noise in the laboratory.

STIMULUS PRESENTATION IN THE LABORATORY. In the laboratory, visual stimuli were projected by an LCD video projector (NEC, model LT157, $1,024 \times 768$ pixels, 60 Hz; maximum luminance 400 cd/m^2) onto a back-projection screen, producing a 52×40 -cm image 1 m away from the observer. Subjects viewed the screen while sitting upright and used a chinrest and forehead rest to maintain stable head position. Auditory stimuli were presented dichotically through high-fidelity headphones (RadioShack Titanium Diaphragm).

STIMULUS PRESENTATION IN THE SCANNER. In the MRI scanner, visual stimuli were projected by the same model LCD video projector (NEC, model LT157, $1,024 \times 768$ pixels, 60 Hz; maximum luminance 400 cd/m^2) onto a back-projection screen, positioned near the subject's waist at the opening of the scanner bore. This produced a 52×40 -cm image 1 m away from the observer. In a subset of scans we measured eye movements while presenting visual stimuli binocularly through Avotec system goggles at a resolution of $1,024 \times 768$. The retinal images produced by the Avotec goggles and the back-projection screen were similar, and we found no significant differences in behavioral or fMRI responses between the two visual presentation systems. Head position in the scanner was stabilized by a bite bar.

Auditory stimuli were presented dichotically through specially designed, magnetically compatible headphones (Resonance Technologies), which attenuate ambient scanner noise by 15–25 dB. Subjects listened to auditory stimuli through earplugs, which attenuate auditory stimulus volume by 30–35 dB. Mean volume levels were adjusted to minimize acoustic artifacts arising from scanner noise (100–115 dB) and to maintain performance at threshold on the auditory task. Each subject was presented with the same mean volume level across scans and volume levels were roughly constant across observers. Thus any differences in performance on our auditory-frequency discrimination task due to differences in sound pressure level should have been minimized. In addition, observers judged the mean volume levels during scanning to be subjectively similar to those used during training, when no scanner noise was present. Importantly, thresholds on the auditory task were not significantly different inside and outside the scanner environment.

Psychophysical methods

Two visual and two auditory stimuli were presented concurrently on each trial. Visual stimuli consisted of drifting sinusoidal gratings of spatial frequency 0.4 cycles/deg, moving over a range of baseline speeds $7^\circ/\text{s}$ ($\pm 25\%$). Gratings were contained within 10° -diameter circular apertures, centered 11° to the left and right of central fixation.

Auditory stimuli consisted of tones (500–1,550 Hz). A different frequency was presented simultaneously to each ear through headphones (the minimum frequency difference between the two ears was ≥ 105 Hz). For this dichotic stimulus, the tone presented to the left ear was perceived as lateralized to the left side of space (left of the central sagittal plane of the head) and the tone presented to the right ear was perceived as lateralized to the right side of space (right of the central sagittal plane of the head). Note that no attempt was made to localize visual and auditory stimuli to the same location in

external space. Rather, auditory stimuli were lateralized to the same side of space as visual stimuli.

Each trial began with a fixation stimulus (250 ms). The shape of the central fixation stimulus endogenously cued observers to attend to one of the four stimuli described earlier. A rightward pointing fixation stimulus cued subjects to direct attention to the right visual stimulus, whereas a leftward pointing fixation stimulus cued attention to the left visual stimulus. For the auditory stimuli, an upward or downward pointing fixation stimulus cued subjects to attend to the left or right ear, respectively. The fixation period was followed by the simultaneous presentation (500 ms) of the two visual and the two auditory stimuli. After a brief delay, where only the fixation stimulus was presented (100 ms), all four stimuli were presented a second time (500 ms). The speed or frequency of each of the four stimuli was presented at baseline in one interval and above baseline in the other interval. The interval with the faster speed or greater frequency was randomly varied from trial to trial, independently for each of the four stimuli. This design minimized anticipation of and habituation to a given stimulus and prevented the use of pairwise comparisons between stimuli in solving the task.

Observers had 700 ms to perform a two-interval, forced-choice discrimination task on the cued stimulus. They indicated, with a key press, whether the first or second presentation of the cued stimulus contained either the faster-moving grating (the visual task) or the higher-frequency tone (the auditory task) (see Fig. 1A). Feedback was provided at the end of each trial: the color of central fixation indicated whether responses were correct, incorrect, or too slow. A new trial started every 2,500 ms regardless of the subject's response. Observers were instructed to maintain gaze at fixation for the duration of the trial.

To ensure stable performance at threshold during scanning, before scanning subjects were trained on the psychophysical task in the laboratory over 3 days (~1,200 trials per subject, ~300 trials per attention condition). During training, the size of the speed or frequency increment was varied from trial to trial using a one-up/three-down staircase (50 trials/staircase). During scanning, we used the speed and auditory frequency increments that supported 80% correct performance. Across observers, the Weber fraction supporting 80% correct performance on the visual speed discrimination task ranged from 9 to 20% and on the auditory task from 0.8 to 4%. In all, 29,952

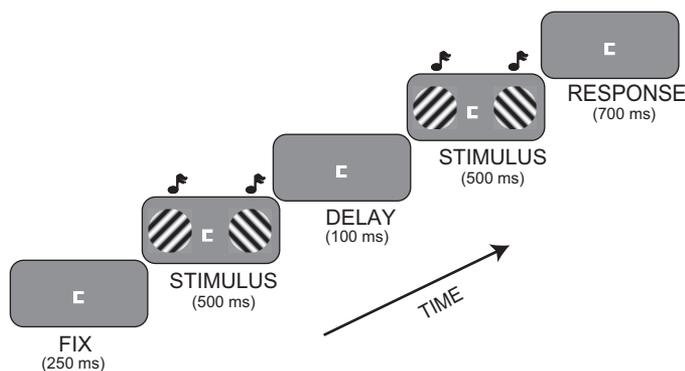


FIG. 1. Stimuli and psychophysical procedure. Sequence of events in a trial: trials began with a central fixation stimulus (fixation interval). After 250 ms, 4 stimuli were presented simultaneously (stimulus period) for 500 ms. There were 2 drifting gratings, 11° to the right and left of fixation and a different auditory tone to each ear. After a 100-ms delay (delay interval) all 4 stimuli were presented again for 500 ms (stimulus interval). Subjects attended one of the 4 stimuli, based on the shape of central fixation. At the attended stimulus, subjects had to judge whether the cued stimulus moved faster or was a higher frequency in the 1st or 2nd stimulus interval. Judgments were made by button press within 700 ms of stimulus offset. Visual feedback was provided at the end of each trial and a new trial was presented every 2,500 ms, regardless of response.

psychophysical trials were completed (1,248 trials per observer per attention condition).

General fMRI methods

Functional MR images for retinotopic mapping and high-resolution T1-weighted anatomical images were acquired on a 1.5-Tesla Siemens Vision system scanner at Thornton Hospital at the University of California, San Diego (UCSD) medical center or on a Signa EXCITE 3-Tesla GE "short-bore" scanner at the Center for Functional Magnetic Resonance Imaging at the UCSD medical school campus. Reference volumes and functional data for our attention conditions were acquired on the GE scanner with a GE eight-channel head coil, using a T2*-weighted echo planar imaging (EPI) pulse sequence (acquisition bandwidth ± 62.5 kHz) lasting 260 s (TR = 2,000 ms, flip angle = 90°). Each scan included 130 acquisitions of 28 4-mm-thick axial slices with 4×4 -mm resolution [field of view (FOV) = 256 mm]. To avoid magnetic saturation effects, the first 10 temporal frames (20 s) of data were discarded from analysis.

At the beginning of each scanning session, localizer scans were collected using the same pulse sequence as cited earlier, lasting 260 s (TR = 2,000 ms, flip angle = 90°). These localizer scans were used to constrain the activation within each of our regions of interest (ROIs) in visual and auditory cortex. At the end of each scanning session, an anatomical scan was collected using a T1-weighted gradient echo pulse sequence (MPRAGE, $1 \times 1 \times 4$ -mm resolution). This lower-resolution anatomical scan allowed us to coregister functional data gathered across many sessions to each observer's high-resolution anatomical scan (an MPRAGE, $1 \times 1 \times 1$ -mm resolution image from the Siemens scanner or an SPGR, $1 \times 1 \times 1$ -mm resolution image from the GE scanner).

fMRI protocol

We used a blocked design where, for each scan, the cue alternated between only two of the four attention conditions we studied. Each scan consisted of 104 trials over 206 s, containing a total of 13 blocks of 8 trials per 20-s block (52 trials for each attention condition). The order of the two attention conditions presented on any given scan was counterbalanced across scans, across days, and across subjects. Eight scans were completed in a given day's scanning session, two repeats of each of the four pairwise comparisons. Each subject participated in six scanning sessions, a total of 12 scans of each of the four pairwise comparisons, for a total of 288 scans across the six observers. We measured eye movements in three subjects in a single scanning session for each subject, consisting of two scans of each of the four pairwise comparisons.

To describe the attention conditions measured with fMRI, we use a nomenclature where the first symbol (*V* or *A*) refers to the brain area measured (visual or auditory cortex) and the second symbol (*V* or *A*) refers to the stimulus subjects are attending (a visual or auditory stimulus). Furthermore, we use a subscript (*c* or *i*) to indicate whether the spatial location of the attended stimulus is contralateral or ipsilateral to the cortical area being measured (see Table 1).

For example, VA_c refers to responses in left visual cortex while subjects perform the auditory task on the contralateral (right) side of space and also refers to responses in right visual cortex while subjects perform the auditory task on the contralateral (left) side of space. Likewise, VA_i refers to responses in left visual cortex while subjects attend auditory stimuli on the ipsilateral (left) side of space, and to responses in right visual cortex while subjects attend auditory stimuli on the right side of space. Similarly, for auditory cortex, AA_c refers to responses in left and right auditory cortex while subjects perform the auditory task on the corresponding contralateral side of space.

Our experimental design had four possible attention conditions: at any given time subjects attended the left visual stimulus, the right visual stimulus, the left auditory stimulus, or the right auditory

TABLE 1. *fMRI responses to four attention conditions both in visual cortical areas and in auditory cortical areas*

Attention Condition	Measured Cortical Hemisphere	Attended Modality	Attended Hemifield
VV_c	Visual cortex: right (<i>left</i>)	Visual	Contralateral: left (<i>right</i>)
VV_i	Visual cortex: left (<i>right</i>)	Visual	Ipsilateral: left (<i>right</i>)
VA_c	Visual cortex: right (<i>left</i>)	Auditory	Contralateral: left (<i>right</i>)
VA_i	Visual cortex: left (<i>right</i>)	Auditory	Ipsilateral: left (<i>right</i>)
AA_c	Auditory cortex: right (<i>left</i>)	Auditory	Contralateral: left (<i>right</i>)
AA_i	Auditory cortex: left (<i>right</i>)	Auditory	Ipsilateral: left (<i>right</i>)
AV_c	Auditory cortex: right (<i>left</i>)	Visual	Contralateral: left (<i>right</i>)
AV_i	Auditory cortex: left (<i>right</i>)	Visual	Ipsilateral: left (<i>right</i>)

For a given attention condition, the first symbol indicates the cortical area whose responses we measured, visual or auditory cortex. The second symbol indicates the modality that was attended, visual or auditory. The subscript indicates the relationship between the brain hemisphere from which responses were measured and the hemifield that was attended, whether the attended stimulus is contralateral or ipsilateral to the cortical hemisphere being measured. For example, condition VA_c measures responses in right visual cortex while attention is directed to an auditory stimulus in the left, or contralateral, hemifield, as well as complementary responses in the left visual cortex (shown in parentheses) while the right auditory stimulus is attended.

stimulus. This allowed us to consider responses in visual cortex to the same visual stimulus: 1) when the contralateral visual stimulus was attended, VV_c ; 2) when it was ignored with attention directed to the ipsilateral visual stimulus, VV_i ; 3) when it was ignored with attention directed to a contralateral auditory stimulus, VA_c ; and 4) when it was ignored with attention directed to an ipsilateral auditory stimulus, VA_i . In a block design, these four states result in six possible pairwise comparisons. Given that each subject was tested repeatedly and that scanning time was limited, we tested only a subset of four of the six possible comparisons. We measured blood oxygen level-dependent (BOLD) responses for the following four pairwise comparisons: $VV_c - VV_i$, $VV_c - VA_c$, $VV_i - VA_i$, and $VA_c - VA_i$. These pairwise comparisons allow us to quantify attention effects between two visual stimuli, two auditory stimuli, or between a visual and an auditory stimulus.

We also measured complementary responses in auditory cortex to the same auditory stimulus: 1) when the auditory stimulus was attended, AA_c ; 2) when it was ignored to attend an auditory stimulus on the ipsilateral side of space, AA_i ; and 3) when it was ignored to attend a visual stimulus on the contralateral, AV_c , or 4) ipsilateral side of space, AV_i . Here we measured BOLD responses to the same auditory stimulus for the following four pairwise comparisons: $AA_c - AA_i$, $AA_c - AV_c$, $AV_i - AA_i$, and $AV_c - AV_i$.

Analysis

PSYCHOPHYSICAL DATA ANALYSIS. For each observer, psychometric functions from the one-up/three-down staircase procedure were fit with a Weibull function using a maximum-likelihood procedure. This Weibull function was used to estimate the visual speed and auditory frequency increments supporting 80% correct performance. These speed and frequency increments were then used during scanning. Performance for each observer was monitored during scanning and average percentage correct performance was computed for each attention condition across scans and across observers.

EYE MOVEMENT ANALYSIS. Eye position was sampled continuously at 60 Hz during scanning. Eye blinks were excluded from analysis for accelerations of the eye that exceeded a threshold of $1^\circ/s^2$. Linear regression was used to remove linear trends and offsets in the data over the course of the entire scan. Once these artifacts were removed, we computed the mean horizontal position of the right eye on a trial-by-trial basis. For each trial, only the 1,000-ms interval when the visual and auditory stimuli were presented was included for analysis. Horizontal eye position for each trial was averaged across all blocks of each attention condition (8 trials/block, 24 blocks/attention condition), for each of the three subjects studied.

fMRI DATA ANALYSIS. *Retinotopic mapping.* Standard retinotopic mapping, cortical segmentation, and cortical flattening techniques

were used to define cortical visual areas V1, V2, V3, V4, and V3A (Boynton et al. 1999; Engel et al. 1994; Sereno et al. 1995). To map retinotopic visual areas, three different sets of stimuli were presented on a mean gray background: expanding rings, rotating wedges, or alternating "hourglass"- and "bowtie"-shaped stimuli. Stimuli were black and white checkerboard patterns that counter-phase flickered at 8 Hz (mean luminance of 340 cd/m² at 100% contrast). The width of the ring stimulus subtended 1/6 of its radius, the polar angle of the wedge stimulus subtended 45°, and the polar angle of the meridian stimulus subtended 90° about the vertical or horizontal meridian.

Area MT+, likely to be the human homologue of MT and MST in the macaque, was selected as a contiguous group of voxels lateral to the parietal-occipital sulcus and beyond the retinotopically organized visual areas, with a time series that correlated ($r > 0.3$ within a <11-s lag time) with the temporal alternation of a moving versus stationary field of dots (Heeger et al. 1999; Huk and Heeger 2002; Tootell and Taylor 1995; Tootell et al. 1995).

Voxels belonging to the cortical surface were identified from the high-resolution anatomical scan using a Bayesian classification algorithm (Teo et al. 1997). Each hemisphere's occipital lobe and temporal lobe were then computationally flattened using a multidimensional scaling algorithm (Engel et al. 1997; software available at <http://white.stanford.edu>). Functional data from a given scanning session were aligned to the high-resolution volume anatomy and then projected onto the flattened cortical representation in three steps: 1) corresponding anatomical locations between the in-plane sections of the functional data and the volume anatomy were marked by eye, 2) the optimal rotation and translation between these two sets of points was determined (Arun et al. 1987), and 3) the measured fMRI BOLD responses were projected onto the flattened representation. For points not containing a functional measurement, data were interpolated using a blurring algorithm that took a weighted average of neighboring pixels containing functional measurements. Following cortical segmentation, flattening, and projection, the boundaries of visual areas were delineated based on the borders between phase transitions (see Fig. 2A for an example of retinotopic maps on an inflated brain image in a single subject).

ROI-based analysis. Regions of activation within each subject's uniquely defined retinotopic ROIs were constrained based on a localizer scan collected at the beginning of each day's scanning session. The localizer scan was presented as a block design with six cycles of 20 s on and 20 s off. During each 20-s "on" period of the localizer scan, two flickering checkerboards covering the same visual angle and eccentricity as the visual stimuli used in the experiment were presented. During each 20-s "off" period, a uniform gray field of the same mean luminance was presented. Voxels within predefined visual areas were considered for analysis if their activity correlated with the presence and absence of the visual stimulus in the localizer scan at a correlation threshold of 0.4 (see Fig. 2, B and C for an example of

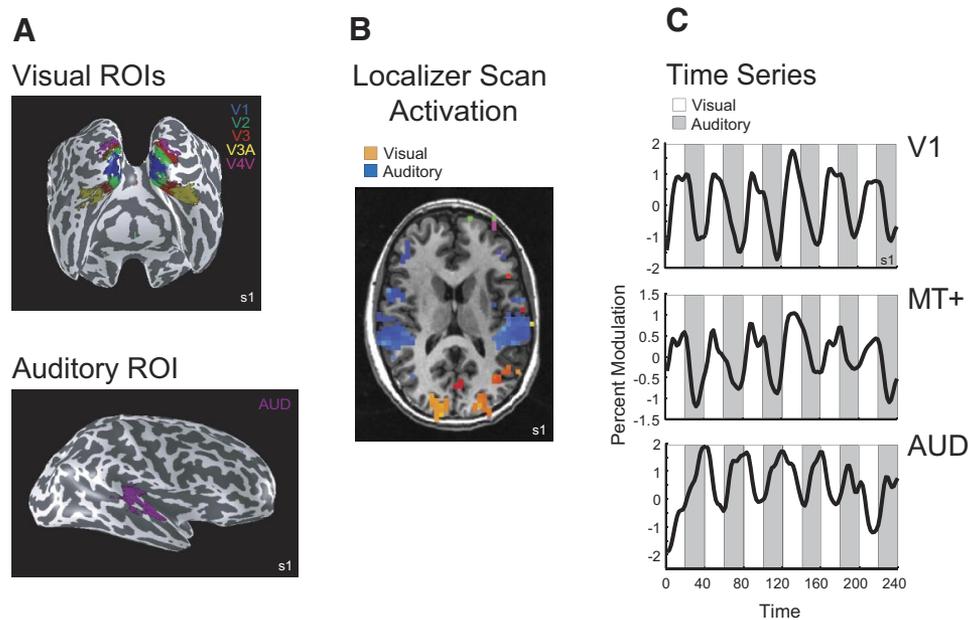


FIG. 2. Regions of interest (ROIs) in early visual and auditory cortex and localizer scan activation. ROIs in early visual cortex, areas V1 (blue), V2 (green), V3 (red), V3A (yellow), and V4V (pink) and early auditory cortex (purple) are shown on an inflated brain for a single subject (A). Brain responses during the localizer scan are shown for this same subject, for a single scan, in a single axial slice (B). Voxels shown in orange reflect activity during visual stimulus presentation (counter-phase modulated flickering checkerboard, lasting 20 s for a total of 6 cycles), whereas voxels shown in blue reflect activity during auditory stimulation presentation (random sequence of auditory tones, same in both ears, lasting 20 s, for a total of 6 cycles). For the axial slice shown here, the average auditory activation on the right was at $z \approx 10$ mm, whereas on the left it was at $z \approx 15$ mm. Functional magnetic resonance imaging (fMRI) time series of the blood oxygen level–dependent (BOLD) response is shown in 3 different ROIs (V1, MT+, and AUD) for this subject, for the same localizer scan, where periods of visual stimulus presentation (white) alternated with periods of auditory stimulus presentation (gray) every 20 s (C).

localizer activation in a single subject in a single axial section and in a given ROI in the left hemisphere over the time course of the localizer scan, respectively). There were no significant differences between hemispheres; thus data were collapsed for analysis across corresponding regions of the left and right visual cortex. Importantly, the measured effects were similar across a range of correlation thresholds commonly used in studies of visual cortex (data not shown).

Auditory areas were defined functionally as well as anatomically. To functionally define an auditory cortical region, we used a reference scan that randomly presented the full range of auditory frequencies to the left ear alone for 20 s and then to the right ear alone for 20 s. A contiguously active region of voxels on a flattened representation centered on Heschl's gyrus was then outlined for each subject. Right auditory cortex included those voxels in phase with left ear stimulation, and left auditory cortex those voxels in phase with right ear stimulation at a correlation threshold of 0.1 (see Fig. 2A for an example of the right auditory ROI on an inflated brain image in a single subject). We used a lower correlation threshold to define auditory areas, compared with that used for visual areas, to compensate for the lower sensitivity and signal-to-noise ratio (SNR) found in auditory areas due to their greater distance from the head coil. Importantly, such a functionally defined auditory region should emphasize auditory areas that are more spatially selective. To anatomically define auditory cortical regions, we followed procedures similar to those of Jancke and colleagues (1999). We defined primary auditory cortex as the region between Heschl's sulcus and the first transverse sulcus in at least two coronal sections and defined secondary auditory cortex as the region lateral and posterior to this area in at least two to three axial sections.

For our functionally defined auditory ROI, the region of activation was constrained based on a localizer scan collected at the beginning of each day's scanning session. The localizer scan was a block design with six cycles of 20 s on and 20 s off. During the 20-s "on" period a random sequence of the same auditory frequencies used in the

experiment was presented dichotically, to the left and right ear simultaneously. During the 20-s "off" period no auditory tones were presented (beyond scanner noise). Voxels active for this localizer scan, within the predefined auditory ROI, were considered for analysis if they were active at a correlation threshold of 0.4 (see Fig. 2, B and C for an example of localizer activation in a single subject in a single axial section and in the auditory ROI in the left hemisphere over the time course of the localizer scan, respectively). Similar effects were observed across a range of correlation thresholds (data not shown).

The Talairach coordinates for the center of activated voxels in our functionally defined auditory ROI (AUD) averaged across our subjects were $x = -48.5 \pm 1.5$, $y = -25.2 \pm 3.1$, $z = 12.5 \pm 2.2$ for the left auditory area and $x = 55.2 \pm 2.2$, $y = -24.2 \pm 3.4$, $z = 10.2 \pm 2.1$ for the right auditory area. These regions of activation correspond to regions within Brodmann area 41 ($x = \pm 25$, $y = -43$, $z = 9$) and area 42/22 ($x = \pm 65$, $y = -25$, $z = 14$), the primary (A1) and secondary (A2) auditory cortex, respectively (see center of activation in Penhune et al. 1996 for primary auditory; Westbury et al. 1999 for secondary auditory). Data were collapsed across corresponding regions of left and right auditory cortex because there were no significant differences between hemispheres in A1 and A2 within each subject across attention conditions (only one of our right-handed subjects showed a significant difference, but only in A2 and only in one of the four attention comparisons, $AA_c - AA_i$).

The Talairach coordinates for the center of our anatomically defined auditory ROIs averaged across subjects for left primary auditory cortex (A1) were $x = -45 \pm 4.8$, $y = -20.5 \pm 3.8$, $z = 8.2 \pm 3.4$ and for right A1 were $x = 46.5 \pm 5.4$, $y = -19.5 \pm 4.5$, $z = 8.3 \pm 1.9$. For secondary auditory cortex (A2), left A2 coordinates were $x = -54.5 \pm 2.8$, $y = -30.2 \pm 3.4$, $z = 9.7 \pm 3.4$ and right A2 coordinates were $x = 54.2 \pm 2.2$, $y = -29.8 \pm 5.6$, $z = 10.5 \pm 2.8$.

Measuring response amplitudes. fMRI responses were quantified by 1) dividing the time series at each voxel by its mean intensity to normalize for differences in the absolute activity range across voxels and provide a percentage change estimate; 2) subtracting linear trends

over time from each voxel's time series to normalize for changes in baseline over time; 3) defining overall signal for a given ROI by averaging the time series across the voxels contained in a given ROI; 4) calculating the phase and amplitude of the sinusoid that best fits the average time series data; 5) representing the phase and amplitude for a given ROI with a single number by computing the component of the vector representing the phase and amplitude of the fMRI signal in phase with the reference stimulus response; and 6) calculating SE of the fMRI response, based on independent measures taken across repeats of the same condition across different scans, rather than within the same scan.

If the voxel in a given brain region is differentially activated by two attention conditions, its image intensity should show periodic cycling over time, associated with the periodic cycling of the two 20-s-long attention conditions. The amplitude of the 40-s cycle fMRI response therefore provides a measure of the amplitude of the difference in the underlying neurophysiological response between the two attention conditions.

Note that our use of a simple block design fMRI experiment where all sensory stimuli are held constant and only attention is varied should minimize any direct contributions from scanner noise on the BOLD response; i.e., the same scanner noise is present irrespective of attention condition, and thus should be subtracted out. Importantly, we are measuring the effects of auditory attention, and not auditory stimulus processing, in which case the use of a sparse sampling design for fMRI data acquisition might be more important.

RESULTS

All observers were well practiced on the task before scanning. Importantly, the Weber fraction supporting 80% correct remained similar inside and outside the scanner for each subject, indicating that subjects were well trained on our task and scanner noise did not have an appreciable influence on performance in our auditory task. During scanning, visual speed and auditory frequency increments were chosen to maintain performance at about 80% correct for both visual and auditory attention conditions. Thus differences in fMRI responses were unlikely to be accounted for by differences in task difficulty because auditory and visual tasks were equally demanding.

Psychophysical performance

First we show psychophysical performance on our four different attention conditions during scanning. Figure 3A plots mean percentage correct performance on each of the attended stimuli, averaged across all six observers. Mean percentage correct performance was 81.1 and 79.6% for the left and right visual stimulus, respectively, and 78.6 and 77.8% for the left and right auditory stimulus, respectively. A repeated-measures ANOVA with subject as a random factor revealed no significant differences in performance across attention conditions ($P > 0.05$). Thus differences in fMRI responses are not likely to be accounted for by differences in performance across our attention conditions.

Differences in fMRI responses were also not due to differences in performance based on the attention comparisons in a given scan (attending between two visual stimuli, two auditory stimuli, or a visual and an auditory stimulus). Our fMRI design measured sustained attention effects and minimized the alternation between attention conditions, with subjects switching between one of two attention conditions, but only at the beginning of each block of eight trials, for a total of 12

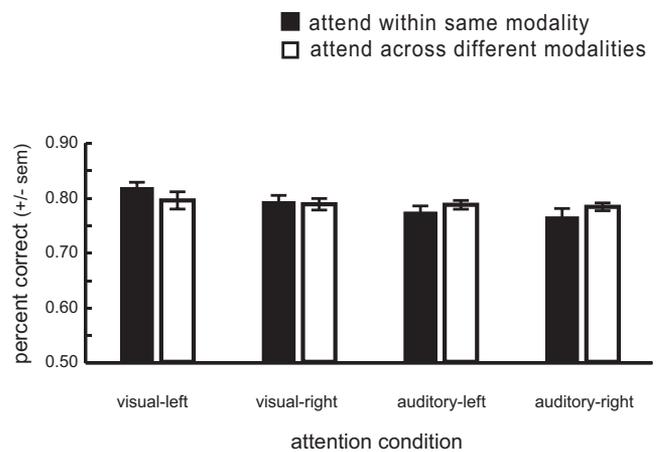


FIG. 3. Psychophysical performance during scanning. Mean percentage correct performance for each of the 4 attended stimuli. Data were averaged across all 6 observers with SE computed across observers. Black bars plot performance at the attended stimulus for blocks of trials where attention was allocated between stimuli within the same modality, across the left and right side of space. White bars plot performance for blocks of trials where attention was allocated between stimuli in different modalities.

switches within a scan. However, there is still the possibility that there might have been some "cost" associated with alternating attention between modalities, or alternatively greater fatigue when subjects alternated attention within a single modality throughout an entire scan. Thus we subdivided our behavioral effects based on whether subjects were allocating their attention within a single modality (white bars of Fig. 3B) or between modalities (black bars of Fig. 3B). We found no significant differences in performance between these conditions for either the visual (two-tailed t -test, $P = 0.2245$) or auditory task (two-tailed t -test, $P = 0.1490$).

Furthermore, differences in fMRI responses were not due to systematic deviations in eye position. We monitored eye position in three of our six subjects during scanning. Mean horizontal eye position across subjects was 0.01° to the left when attending the right visual stimulus, 0.045° to the left when attending the left visual stimulus, 0.015° to the left when attending the right auditory stimulus, and 0.015° to the left when attending the left auditory stimulus. We found no significant differences in mean eye position across our attention conditions [$F(3,18) = 0.4803$, $P = 0.4803$, data not shown].

fMRI results

Again, our goal was to compare the magnitude of spatial and cross-modal attention and determine how they interact across early visual and auditory cortical areas. As described earlier, our design had four attention conditions (attend right visual, left visual, right auditory, or left auditory). We measured four pairwise comparisons between attention conditions in visual cortex: $VV_c - VV_i$, $VV_c - VA_c$, $VV_i - VA_i$, and $VA_c - VA_i$ and the analogous four pairwise comparisons in auditory cortex: $AA_c - AA_i$, $AA_c - AV_c$, $AV_i - AA_i$, and $AV_c - AV_i$.

fMRI responses within visual cortex

Figure 4A shows percentage signal change in the BOLD response for a single example subject. Figure 4B shows the average response across all six subjects. BOLD responses were averaged

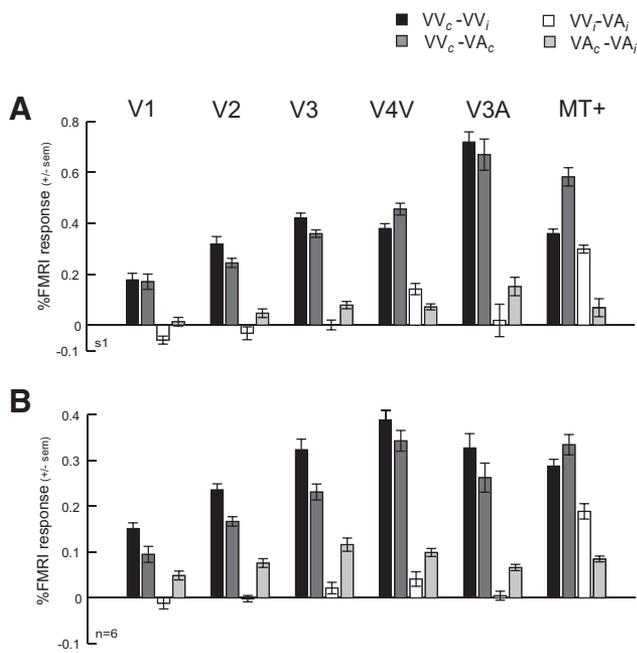


FIG. 4. fMRI responses in early visual cortex. Percentage signal change in the fMRI BOLD response in V1, V2, V3, V4V, V3A, and MT+ for a single subject (A) and across all 6 subjects (B). Data were averaged across the left and right hemispheres and SE was computed across, not within, scans. $VV_c - VV_i$ is the response to a contralateral visual stimulus when it is attended vs. when attention is directed away to a competing ipsilateral visual stimulus (black). $VV_c - VA_c$ is the response to a contralateral visual stimulus when it is attended vs. when attention is directed away to a contralateral auditory stimulus on the same side of space (dark gray). $VV_i - VA_i$ is the response to a contralateral visual stimulus when attention is directed away to an ipsilateral visual stimulus vs. an ipsilateral auditory stimulus (white). $VA_c - VA_i$ is the response to a contralateral visual stimulus when attention is directed away to a contralateral auditory stimulus on the same side of space vs. an ipsilateral auditory stimulus on the opposite side of space (light gray).

across the right and left hemispheres of each visual area for each attention comparison. For each of the early visual areas considered, the attention comparisons we studied were significantly different from zero (across subjects, two-tailed t -test, $P < 0.01$), except $VV_i - VA_i$ in V1, V2, V3, and V3A.

Black bars represent $VV_c - VV_i$, which measures a basic left-right spatial attention effect, as found previously within the visual modality (Gandhi et al. 1999; Martinez et al. 1999; Somers et al. 1999), here considered in the presence of two lateralized auditory distractors. A repeated-measures ANOVA with subject as a random factor indicated that this spatial attention effect differed significantly across visual areas [$F(5,25) = 2.9453$, $P = 0.0318$]. Post hoc tests revealed that the effect of spatial attention was weakest in area V1, replicating findings from electrophysiological recordings in nonhuman primates and functional imaging in humans (reviewed in Treue 2001). Area V4V showed the largest spatial attention effect (Tukey HSD, $P < 0.05$). Furthermore, spatial attention effects in MT+ were slightly, but not significantly, smaller than those in V4V and V3A. A slight reduction in the magnitude of our spatial attention effect is not particularly surprising: human MT+ includes MST (medial superior temporal), an area known to have large receptive fields extending into the ipsilateral visual field, which might reduce differences between attending to ipsilateral versus contralateral stimuli.

Dark gray bars represent $VV_c - VA_c$, the response to a visual stimulus when it is attended compared with when it is ignored to attend an auditory stimulus on the same side of space. Positive values indicate stronger responses in this condition when the visual stimulus is attended. There was again a significant effect of visual area [ANOVA, $F(5,25) = 5.4004$, $P = 0.0017$]. If we think of attention as decreasing responses to irrelevant stimuli, rather than simply enhancing responses to relevant stimuli, then this is a measure of the extent to which information is allowed to pass through early sensory areas when subjects attend another modality. The reduced response when attending another modality was significantly smaller in V1 than in V4V, with no significant differences between MT+, V4V, and V3A (Tukey HSD, $P < 0.05$). Thus in V1 this cross-modal effect ($VV_c - VA_c$) was weaker than the basic spatial attention effect ($VV_c - VV_i$). In contrast, a more central visual area, MT+, had cross-modal effects that were larger than spatial attention effects (Tukey HSD, $P < 0.05$).

White bars represent $VV_i - VA_i$, the difference in the response to an ignored visual stimulus when subjects attend to an ipsilateral visual stimulus compared with an ipsilateral auditory stimulus. Positive values here indicate stronger responses to the same unattended visual stimulus when redirecting attention to another visual compared with an auditory stimulus. Overall, this cross-modal comparison produced a fairly weak modulation that showed significant differences across visual areas [ANOVA, $F(5,25) = 5.0976$, $P = 0.0023$]. Here, the largest cross-modal effects were observed in MT+, which showed weaker responses to an ignored visual stimulus when subjects attended an auditory stimulus compared with when they attended a visual stimulus. Post hoc tests revealed a significant difference between MT+ and all other visual areas (Tukey HSD, $P < 0.05$).

Finally, light gray bars represent $VA_c - VA_i$, the response to an ignored visual stimulus, when subjects alternate attention between a contralateral and an ipsilateral auditory stimulus. All visual areas show a positive value for this attention comparison, indicating larger responses when attending to a stimulus in a different modality that shares the same side of space as the ignored visual stimulus whose fMRI responses are being measured. There were no significant differences across visual areas for this comparison [ANOVA, $F(5,25) = 1.0668$, $P = 0.4021$].

fMRI responses within auditory cortex

To determine whether our basic pattern of results could be generalized to another early sensory cortical area, we performed a complementary analysis of our four pairwise attentional comparisons in auditory cortex. As described earlier, the average Talairach coordinates for the functionally defined auditory areas we isolated in our reference scan (AUD) coincide with the location of primary/secondary auditory cortex. In addition, we examined effects within subdivisions of auditory cortex, primary (A1) and secondary (A2) auditory cortex, as defined by anatomical landmarks.

Percentage signal change in the BOLD response in our auditory ROIs is plotted for the single example subject (Fig. 5A), as well as for the average across all six subjects (Fig. 5B). BOLD responses were averaged across the right and left hemispheres of each auditory area for each attention comparison. Each of the early auditory areas considered showed

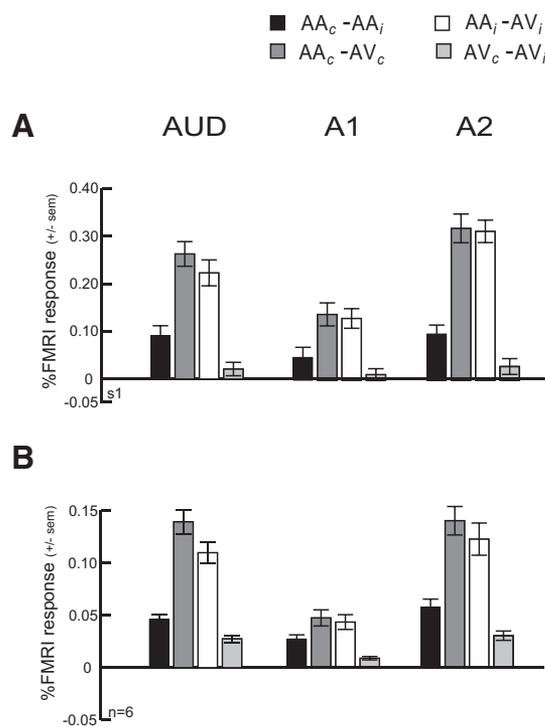


FIG. 5. fMRI responses in early auditory cortical areas. Percentage signal change in the fMRI BOLD response in functionally defined auditory cortex (combined across primary and secondary auditory areas) and in anatomically defined primary (A1) and secondary (A2) auditory cortex for a single subject (A) and across all 6 subjects (B). Data were averaged across the left and right hemispheres and SE was computed across, not within, scans. $AA_c - AA_i$ is the response to a contralateral auditory stimulus when it is attended vs. when attention is directed away to a competing ipsilateral auditory stimulus on the opposite side of space (black). $AA_c - AV_c$ is the response to a contralateral auditory stimulus when it is attended vs. when attention is directed away to a contralateral visual stimulus on the same side of space (dark gray). $AA_i - AV_i$ is the response to a contralateral auditory stimulus when attention is directed away to an ipsilateral auditory stimulus vs. an ipsilateral visual stimulus (white). $AV_c - AV_i$ is the response to a contralateral auditory stimulus when attention is directed away to a contralateral visual stimulus on the same side of space vs. an ipsilateral visual stimulus on the opposite side of space (light gray).

attention effects that were significantly different from zero (across subjects, one-tailed t -test, $P < 0.05$), except $AV_c - AV_i$ in A1 (one-tailed t -test, $P = 0.0535$).

Black bars show a basic left–right spatial attention effect within the auditory modality, $AA_c - AA_i$, for all auditory ROIs analyzed. A repeated-measures ANOVA with subject as a random factor revealed no significant differences across auditory areas for this attention comparison [$F(2,10) = 3.6565$, $P = 0.0643$]. As in earlier studies (e.g., Alho et al. 1999; Jancke et al. 2003; Lipschutz et al. 2002; Petkov et al. 2004; Woldorff et al. 1993) we found auditory spatial attention effects with dichotic listening, here in the presence of visual distractors.

The dark gray bars show the $AA_c - AV_c$ comparison, with positive responses indicating a larger response to an auditory stimulus when it is attended compared with when it is ignored to attend a contralateral visual stimulus. Here we found a significant effect of auditory area [$ANOVA$, $F(2,10) = 15.5148$, $P = 0.0009$]. Post hoc tests revealed that this cross-modal effect was significantly weaker in A1 (Tukey HSD, $P < 0.05$). Interestingly, such cross-modal effects were almost

twofold larger than the effects of spatial attention in auditory cortex. This is a much larger effect than that observed in visual cortex, except in area MT+ where cross-modal attention effects were larger than spatial attention effects.

White bars show the $AA_i - AV_i$ comparison, the response to an ignored auditory stimulus when attention is directed to an ipsilateral auditory stimulus versus an ipsilateral visual stimulus. This cross-modal attention effect was positive, suggesting that responses to unattended auditory information were stronger when subjects attended an auditory stimulus than when they attended a visual stimulus. There was a significant effect of auditory area [$ANOVA$, $F(2,10) = 9.0500$, $P = 0.0057$], with post hoc tests finding significantly weaker effects in A1 (Tukey HSD, $P < 0.05$).

Finally, the light gray bars represent $AV_c - AV_i$, the response to an ignored auditory stimulus when attention is directed to a contralateral versus an ipsilateral visual stimulus. This comparison, of the effects of spatial attention across modalities, produced the weakest modulation in our auditory ROIs. Responses to an ignored auditory stimulus were slightly stronger when the attended stimulus in a different modality, a visual stimulus, shared the same side of space as the ignored auditory stimulus. Similar to results obtained in early visual areas—we found no significant differences in this measure of cross-modal spatial attention across the early auditory areas we considered [$ANOVA$, $F(2,10) = 3.6565$, $P = 0.0643$].

Importantly, although two of our six subjects were left-handed, there were no significant differences in responses in our auditory ROIs as a function of handedness, for the low-level auditory stimuli used in our study (post hoc Tukey HSD, $P > 0.05$).

In sum, we find significant differences between our attention comparisons in both early visual and auditory cortical areas. Note that our experimental design and the results we observed cannot be brought to bear on questions regarding the influence of events within a trial on neuronal responses. For example, we cannot assess the contributions of descending influences from cue-related activity: we used a visual cue to endogenously redirect attention regardless of modality and the cue was a spatially appropriate pointer in the visual modality (pointing left or right for left and right visual stimuli, respectively), but not in the auditory modality (pointing up and down for left and right auditory stimuli, respectively). However, all subjects were well trained and likely to have internalized the meaning of the cue. Furthermore, subjects did not have to react to the novelty of the cue on a trial-by-trial basis because a given cue was presented for eight sequential trials and a given scan alternated only between two cues. Nor can we assess the contribution of feedback-related activity. Although subjects were provided with feedback on a trial-by-trial basis any influence of feedback should have been held constant given that subjects performed at threshold across attention conditions.

Furthermore, our design cannot isolate the unique contributions of shifting attention between modalities or across space from those of maintaining attention at a given modality or region of space. Rather, our results should be dominated by the effects of maintaining attention at a spatial location or modality, given that subjects shifted their attention only after eight trials, or 20 s. Finally, the current experiment cannot distinguish between an attentional mechanism that acts by enhancing

responses to relevant stimuli as opposed to depressing responses to irrelevant stimuli—either mechanism could account for the observed results.

Parameterizing the effects of spatial attention, cross-modal attention, and cross-modal spatial attention

We can summarize the preceding results by describing separate parameters representing the effects of spatial attention, cross-modal attention, and cross-modal spatial attention (the effects of spatial attention on the other modality). Parameterizing our data provides another way to conceptualize our results, rather than a neuronal mechanism to explain our results. We assume that the effects of spatial attention and cross-modal attention each have an additive influence on the fMRI response to a visual stimulus, and that they do not interact with each other. Such an assumption is unlikely to be true in detail, but, as will be shown, is an effective first approximation.

To predict the effects of *cross-modal attention*, we assume that responses to any stimulus in the attended modality, irrespective of spatial location, are enhanced by an additive term α , whereas attention to stimuli in an unattended modality is suppressed by $-\alpha$. We predict *spatial attention* effects by assuming that attending to a stimulus enhances responses to that stimulus by an additive term β , and diminishes responses to any stimulus in the same modality but a different spatial location by $-\beta$. To model *cross-modal spatial attention* (the response to an ignored stimulus in one modality as attention alternates between contralateral and ipsilateral stimuli in a different modality), we assume that attending a stimulus on one

side enhances responses to contralateral stimuli in a different modality by an additive term γ , and diminishes responses to stimuli that are both ipsilateral and in a different modality by $-\gamma$ (see Fig. 6 for a schematic). As an oversimplification, unlikely to hold true in detail, we assume that suppressive and enhance effects are of equal magnitude.

Our three parameters are α , representing cross-modal attention; β , representing spatial attention; and γ , representing the effects of spatial attention on the other modality. For example, if the effects of spatial location are entirely limited to stimuli in the same modality as the attended stimulus, then γ will be 0. On the other hand, if the effects of spatial location operate independently of whether stimuli are in the same modality (i.e., spatial attention effects are the same irrespective of whether attended and ignored stimuli are in the same or different modalities), then γ will equal β . We used a least-squares approximation to solve for our three parameters across the six subjects using the equations subsequently outlined.

Responses in visual cortex for our four attention comparisons are predicted by the following equations

$$\begin{aligned}
 VV_c - VV_i &= (\alpha + \beta) - (\alpha - \beta) = 2\beta \\
 VV_c - VA_c &= (\alpha + \beta) - (-\alpha + \gamma) = 2\alpha + \beta - \gamma \\
 VV_i - VA_i &= (\alpha - \beta) - (-\alpha - \gamma) = 2\alpha - \beta + \gamma \\
 VA_c - VA_i &= (-\alpha + \gamma) - (-\alpha - \gamma) = 2\gamma
 \end{aligned}$$

Analogous responses in auditory cortex are modeled as follows

$$\begin{aligned}
 AA_c - AA_i &= 2\beta \\
 AA_c - AV_c &= 2\alpha + \beta - \gamma
 \end{aligned}$$

gain factors

- α (cross-modal)
- β (spatial)
- γ (cross-modal spatial)

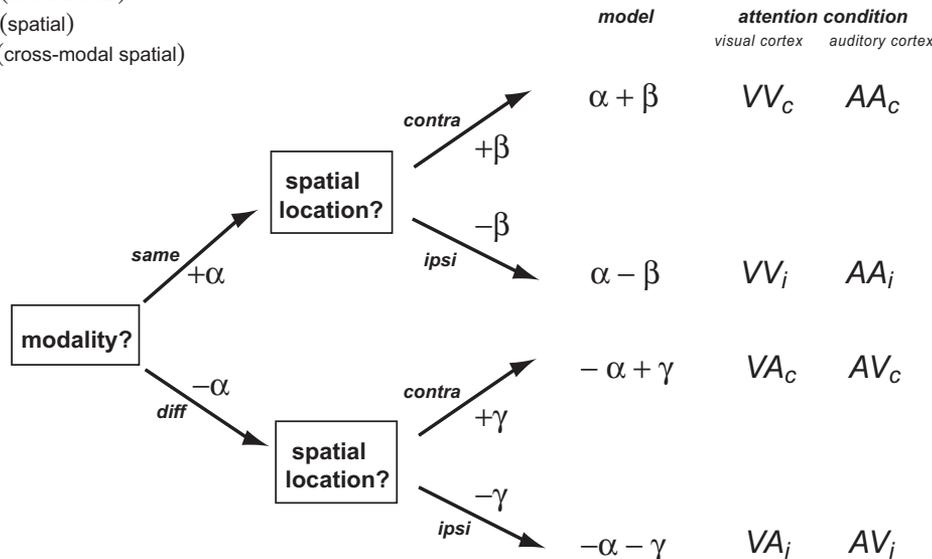


FIG. 6. Schematic diagram of the parameters associated with spatial, cross-modal, and cross-modal spatial attention. Cross-modal effects are modeled by assuming that responses to an attended stimulus in the same modality as the cortical area under study (a visual stimulus if the cortical area is visual) are enhanced by an additive factor α , whereas responses to an unattended stimulus are reduced by $-\alpha$. Spatial attention effects are modeled by assuming that attending a stimulus contralateral to the cortical area under study enhances responses to that stimulus by an additive factor β , and diminishes responses to any stimulus in the same modality but a different spatial location, in our case an ipsilateral location, by $-\beta$. Cross-modal spatial effects, the response to a stimulus in one modality as attention alternates between contralateral and ipsilateral spatial locations and between modalities, are modeled by assuming that attending a stimulus enhances responses to contralateral stimuli in a different modality by an additive factor γ , and diminishes responses to stimuli that are both ipsilateral and in a different modality by $-\gamma$. Thus attending a contralateral stimulus in the same modality as the cortical area studied can be represented as $\alpha + \beta$, our VV_c and AA_c attention conditions in visual and auditory cortex, respectively. Likewise, attending a contralateral stimulus in a different modality can be represented as $-\alpha + \gamma$, our VA_c or AV_c attention conditions. A similar convention holds for the other attention conditions.

$$AA_i - AV_i = 2\alpha - \beta + \gamma$$

$$AV_c - AV_i = 2\gamma$$

Figure 7 shows data (solid bar plots) and model predictions (striped bar plots) for primary visual cortex. Model predictions closely match the data, which is to be expected given that we have three parameters to predict four data points. Fits are similarly good for the other visual areas and for auditory cortical areas (data and predictions not shown).

Figure 8A plots the factors α (cross-modal attention, black squares), β (spatial attention, white circles), and γ (cross-modal spatial attention, gray triangles), across early visual and auditory areas for the single example subject shown earlier. Figure 8B plots these factors averaged across all six subjects.

The results shown here assume that both spatial attention and cross-modal attention factors each have an additive influence on the fMRI response. Electrophysiological experiments demonstrate that the effects of spatial attention can act by a contrast gain, which would be identical to a multiplicative gain for the experiments described herein because they were carried out at a single contrast (Reynolds et al. 2000). However, in fMRI experiments the effects of spatial attention seem to be best described by an additive, rather than a multiplicative, model (Buraças and Boynton 2007; but see also Williford and Maunsell 2006 for an example from physiology experiments in area V4).

When we modeled our effects using an analogous multiplicative model, instead of the additive model shown earlier, we obtained similar results in terms of the relative strength of our spatial, cross-modal, and cross-modal spatial attention parameters. Future studies will need to determine a model that best

describes these types of cross-modal and cross-modal spatial attention effects.

Parameters in visual cortex

All parameters were significantly >0 in all visual areas (one-tailed t -test, $P < 0.05$). Modulation by cross-modal attention, as modeled by α , increased across visual areas, with the weakest effects in V1 and the strongest effects in MT+ [$F(5,25) = 5.9805$, $P = 0.0009$]. The effect of spatial attention, as modeled by β , was weakest in V1 and strongest in V4V [$F(5,25) = 3.5631$, $P = 0.0144$]. A post hoc test revealed no significant differences across V4V, V3, V3A, or MT+ (Tukey HSD, $P > 0.05$). Furthermore, in V1, the spatial attention term was significantly greater than the cross-modal attention term (one-tailed t -test, $P = 0.0230$). However, by MT+ the cross-modal and spatial attention terms were equally strong (one-tailed t -test, $P = 0.5865$).

Finally, our term γ , describing the effects of spatial attention on another modality, was >0 but $<\beta$ in all visual areas. This indicates that the response to an ignored visual stimulus was enhanced when the attended auditory stimulus was on the same side of space compared with the opposite side of space. This cross-modal spatial attention term was smaller than the spatial attention term within the visual modality. A repeated-measures ANOVA with subject as a random factor revealed no significant differences in γ across visual areas [$F(5,25) = 1.1734$, $P = 0.3499$].

Parameters in auditory cortex

All parameters in early auditory areas were significantly >0 (one-tailed t -test, $P < 0.05$). In auditory cortex, a post hoc test (Tukey HSD, $P < 0.05$) revealed that modulation by cross-modal attention, as modeled by α , was significantly weaker in A1 [ANOVA, $F(2,10) = 12.6583$, $P = 0.0018$], as was modulation by spatial attention, as modeled by β [$F(2,10) = 4.8803$, $P = 0.0332$]. Furthermore, our cross-modal attention term showed a trend to being larger than our spatial attention term in A2 ($P = 0.0686$), comparable to results in visual area MT+. No such trend was observed in A1 ($P = 0.1439$).

Our cross-modal spatial attention term γ showed no significant differences across early auditory areas [$F(2,10) = 2.8718$, $P = 0.1034$], similar to results across early visual areas. Thus even the early areas we considered in auditory cortex exhibit robust cross-modal attention effects. These early auditory areas may be considered more analogous to higher-level areas in the visual system because a substantial amount of processing is done subcortically in the auditory modality (e.g., Masterton 1992). Thus the larger cross-modal term observed in A2, which is more comparable to results in MT+, may not be surprising.

In sum, all our parameters were weaker in auditory cortex than in visual cortex. The maximal average of our parameters across subjects in auditory cortex was <0.06 compared with >0.15 in visual cortex.

DISCUSSION

Many studies of cross-modal attention have focused on the higher-level brain areas involved in generating a unified cross-modal percept of the world, when the information across

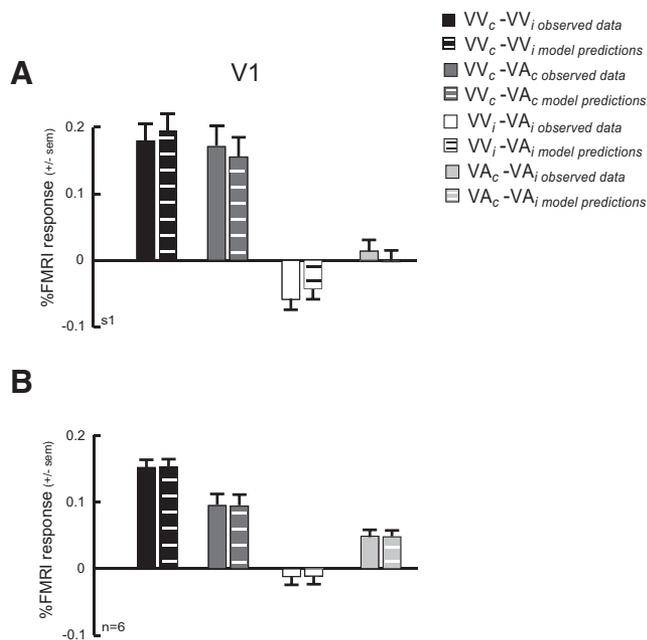


FIG. 7. Predicted fMRI responses. We used our model to predict percentage signal change in the fMRI BOLD response in early visual and auditory cortex. Data (solid fill) and the model predictions (striped fill) are shown for each attention condition, $VV_c - VV_i$ (black), $VV_c - VA_c$ (dark gray), $VV_i - VA_i$ (white), and $VA_c - VA_i$ (light gray), in area V1 for a single subject (A) and across all 6 subjects (B). Data were averaged across the left and right hemispheres and SE was computed across, not within, scans.

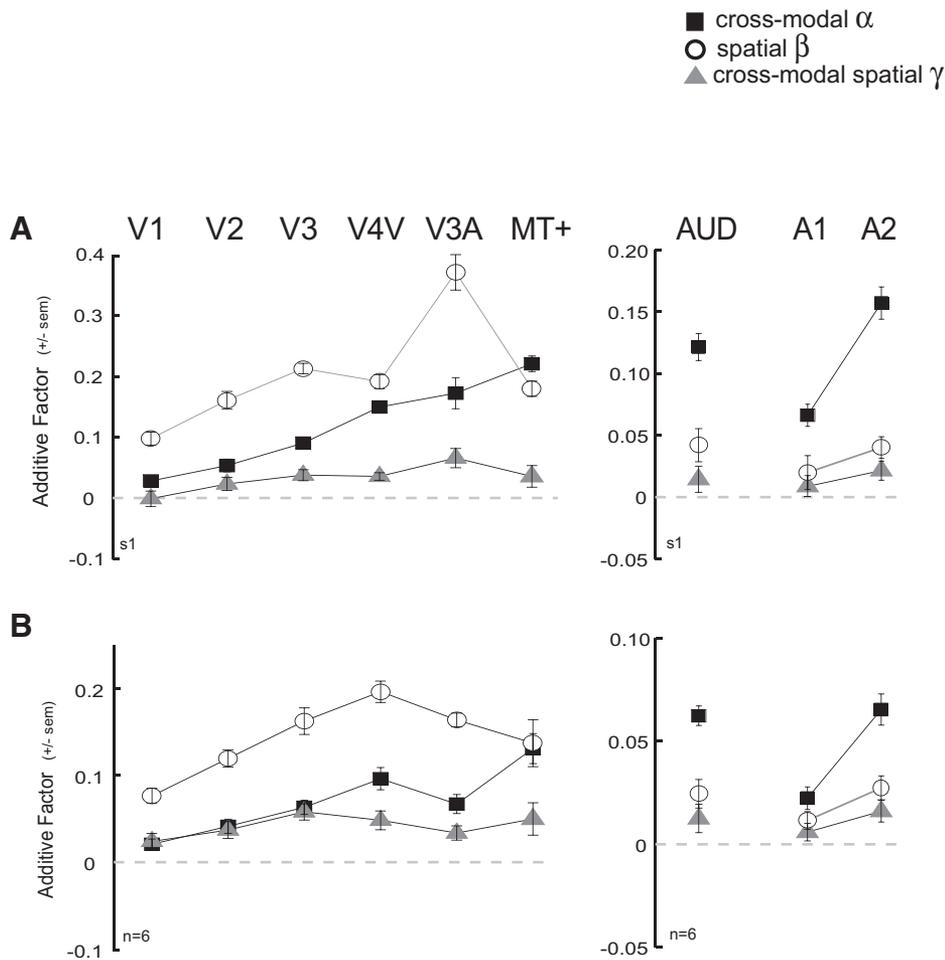


FIG. 8. Estimates of the parameters associated with spatial, cross-modal, and cross-modal spatial attention. Given the 4 pairwise attention comparisons of our experiment a least-squares solution yielded 3 parameters. Estimated cross-modal, α (black squares); spatial, β (white circles); and cross-modal spatial, γ (white triangles) parameters are shown for a single subject (A) and across all 6 subjects (B) for early visual areas. Comparable terms are also shown for a functionally defined auditory cortical area (AUD) and anatomically defined primary (A1) and secondary (A2) auditory cortex. SE was computed across subjects for each cortical area and was calculated using a bootstrapping procedure that resampled the original data over 1,000 repetitions with replacement.

sensory modalities is correlated and integrated. However, in our everyday experience we often need to selectively process information in one modality while ignoring information in a different modality. Such cross-modal influences might be particularly important in early sensory areas not involved in cross-modal integration per se. In fact, recent evidence highlights the presence of cross-modal influences in early sensory areas, areas traditionally considered to be responsive to information in only a single modality.

Here we used stimuli that were not optimized for cross-modal integration and investigated how neuronal responses to simple visual and auditory stimuli vary across early sensory areas and how they depend on the modality and spatial location that subjects attend. Our design allowed us to examine the effects of attention while sensory stimuli, motor responses, eye movements, and task difficulty were held constant, and habituation and anticipation effects were minimized. We find that attending to an auditory stimulus influences responses in visual cortex as early as V1 and can influence responses differently across early visual areas. Similarly, we find that attending to a visual stimulus influences responses in auditory cortex as early as A1. We quantified the relative effects of spatial attention (redirecting attention within the same modality, between a stimulus on the left and right side of space), cross-modal attention (redirecting attention between two different modalities), and cross-modal spatial attention (the effects of spatial

attention on another modality), by deriving values for three parameters, β , α , and γ , respectively.

Spatial attention

Consistent with previous studies (Gandhi et al. 1999; Martinez et al. 1999; Somers et al. 1999; Treue 2001), the effects of spatial attention were smallest in V1, slightly larger in V2, and largest in areas V3 and V4V. Spatial attention effects were smaller, although not significantly smaller, in V3A and MT+ (for a review see Treue 2001).

It is known that receptive field sizes increase across visual areas in both monkeys (e.g., Desimone and Gross 1979;Gattass et al. 1981, 1988; Van Essen et al. 1984) and humans. In humans, estimated MT+ receptive field size is threefold larger than V1 for a given eccentricity (e.g., Kastner et al. 2001). One predicted consequence of larger receptive fields in MT+ would be decreased fMRI modulations when alternating attention between the left and right side of space because ipsilateral as well as contralateral visual stimuli would drive neural responses. This could possibly explain why spatial attention effects were weaker in area MT+. However, we did not find significant differences between attention effects in area MT+ and those in areas with smaller receptive fields, such as V3A, V4V, or V3. Furthermore, our data are also consistent with the idea that more central visual areas may provide little or no additional contribution to spatial attention: the process of

“focusing” the spatial spotlight of attention might occur within lower visual areas with smaller receptive fields.

We also find evidence for spatial attention effects in the auditory domain, consistent with previous findings (e.g., Alho et al. 1999; Lipschutz et al. 2002). Interestingly, in the auditory domain, spatial attention effects were found as early as A1, but showed no significant difference between primary and secondary auditory cortices. Such a result might be expected, given the relatively weak lateralization of spatial location in auditory cortex. Auditory cortex shows a weaker symmetry in the representation of contralateral space than visual cortex and, unlike visual cortex, shows hemispheric specialization for processing certain types of auditory information in the left or right ear and the left or right auditory hemisphere (reviewed in Tervaniemi and Hugdahl 2003; but see Boemio et al. 2005).

One caveat is that because we used headphones to deliver our auditory stimuli in the fMRI environment, subjects may have been switching between extrapersonal and peripersonal space when switching attention between visual and auditory stimuli (a confound pointed out by Spence and Driver 1997). Such a limitation is intrinsic to any fMRI studies using headphones and cannot be excluded as a possible factor in our results. Another caveat is that our spatial manipulation redirects attention across and not within hemifields. However, the results we observe, subsequently described in more detail, are similar to those obtained by Eimer and colleagues (2004) in a task manipulating spatial attention within, rather than across, hemifields. Thus our effects may apply to spatial effects in general, rather than specifically to manipulations across hemifields. Future studies will need to examine spatial factors in greater detail.

Cross-modal attention

Previous work has shown that cross-modal attention can enhance cortical responses to congruent stimuli in two different modalities, potentially accounting for improvements in speed or accuracy on detection and discrimination tasks (e.g., Calvert et al. 2000; Macaluso et al. 2000; Stein 1998; Teder-Saljarvi et al. 2002; for a review see Driver and Spence 2000). Cross-modal attention has also been shown to decrease cortical responses to distracting or competing stimuli (e.g., Foxe et al. 2005; Johnson and Zatorre 2005; Laurienti et al. 2002; Weissman et al. 2004), which could underlie improved performance by diminishing the influence of distracting stimuli in another modality.

In terms of determining cross-modal influences in early sensory areas, most previous studies of cross-modal attention have either focused on a single visual area, only found effects in certain areas (e.g., lingual gyrus: Macaluso et al. 2000; lateral occipital sulcus: MT+, Lewis et al. 2000; MT+, Berman and Colby 2002; MT+, V3a, and kinetic occipital area, Rees et al. 2001), or collapsed across areas (e.g., Laurienti et al. 2002; collapsed across areas 17, 18, 19, 30, 37, cuneus, lingual, and fusiform cortex). We used very simple visual and auditory stimuli and examined effects across all early visual areas, where each area was uniquely defined for each of our subjects in a separate series of scans.

We found that area V1, despite being traditionally considered unimodal, was influenced by cross-modal attention. This was not a general effect of attending elsewhere. Responses to

ignored visual stimuli in V1 were weaker when subjects attended another visual stimulus than when they attended an auditory stimulus, our $VV_i - VA_i$ condition. Furthermore, the effect of attending an auditory stimulus depended on whether it shared the same side of space as the ignored visual stimulus, as subsequently described.

Within V1, spatial attention effects were larger than cross-modal effects: the response to an *attended* visual stimulus in V1 was stronger when attention was redirected to another visual stimulus ($VV_c - VV_i$) than to an auditory stimulus ($VV_c - VA_c$). If we think of attention as suppressing irrelevant responses, rather than enhancing relevant responses, our results for cross-modal effects when a visual stimulus was always ignored ($VV_i - VA_i$) suggest that more information at an ignored visual location is allowed to pass through area V1 when attending an auditory stimulus versus another visual stimulus. The current fMRI results cannot distinguish between these two possible mechanisms of enhancement of relevant stimuli or suppression of irrelevant stimuli.

The cross-modal attention effects we measured varied across the reported visual areas. In MT+, for an *attended* visual stimulus the effects of cross-modal attention ($VV_c - VA_c$) were larger than the effects of spatial attention ($VV_c - VV_i$). Furthermore, for a visual stimulus that was always ignored ($VV_i - VA_i$), responses were larger when this stimulus was ignored to attend another visual stimulus compared with an auditory stimulus. The gradual and systematic change we found across visual areas in the relative contributions of within-modal spatial attention compared with cross-modal attention can be clearly seen in the changes of our estimated additive factors, β and α (see Fig. 7).

Similarly, in auditory cortex cross-modal attention effects were found as early as A1 and were significantly larger in secondary than in primary auditory cortex. Interestingly, as early as A1, cross-modal attention effects ($AA_c - AV_c$) were larger than spatial attention effects ($AA_c - AA_i$), similar to results in MT+ rather than V1. Such a result might be expected, given the relatively weaker lateralization of spatial location in auditory cortex, diminishing the influence of spatial attention. Even for an auditory stimulus that was always ignored ($AA_i - AV_i$), stronger responses were observed when the stimulus was ignored to attend another auditory compared with another visual stimulus. The relative contributions of within-modal spatial attention compared with cross-modal attention in early auditory areas are clearly seen in the estimated factors, β and α (see Fig. 7). Furthermore, in our functionally defined ROI (including both primary and secondary auditory cortices) results were similar to those obtained in anatomically defined ROIs. We might have obtained even cleaner results if, instead of using anatomical landmark-based localizations of functional areas where there is high intersubject variability, we had been able to subdivide auditory areas functionally and thus uniquely for each subject.

The anatomical substrate of cross-modal attention

The differences we observe in attentional effects across visual areas could reflect important limitations imposed by anatomical projections. One possibility is that the anatomical substrate for the cross-modal influences we observe arises from projections to visual cortex directly from auditory cortex.

Direct monosynaptic projections from primary auditory cortex to both V1 and V2 have been found in adult monkeys. These auditory projections are denser in V2 than in V1, are found in portions of V1 and V2 representing the more peripheral and lower visual field (Rockland and Ojima 2003), and their density increases as a function of retinal eccentricity in area V1 (Falchier et al. 2002). It is unknown whether such projections are primarily excitatory or inhibitory or how their density may vary beyond area V2. Thus it is unclear whether such direct auditory connections could influence MT+. Interestingly, behavioral evidence suggests a potential functional role: as predicted by the anatomy, the influences of auditory stimuli on the perception of a sound-induced visual illusion are strongest in the peripheral and lower portions of the visual field (Shams et al. 2002). Our future work will investigate the effects of stimulus location (eccentricity and upper vs. lower visual field) on cross-modal attention to determine whether these direct auditory projections can modulate the effects of cross-modal attention.

Alternatively, anatomical influences could originate from feedback projections, from higher-level areas such as parietal and prefrontal cortex (Cavada and Goldman-Rakic 1989a,b; Lewis and Van Essen 2000). Such areas have already been shown to control the allocation of attention irrespective of modality (Eimer et al. 2003; Shomstein and Yantis 2004; but see Chambers et al. 2004). Furthermore, such higher-level areas receive input from auditory cortex and have strong projections to MT+. Changes in the density of such feedback projections across early visual areas have not been fully characterized. Thus it is unknown how they might contribute to the differences we observe across early visual areas. For example, if these feedback projections are sparser in V1 or V2 than in V3A, V4V, or MT+, they could subserve some of the effects we observe.

A complementary consideration of anatomical influences that could underlie differences between cross-modal attention effects in primary and secondary auditory cortices would be equally important, but is beyond the scope of this paper.

Cross-modal spatial attention

Unlike our other cross-modal effects, which varied across visual as well as auditory areas, there were no significant differences across visual areas for our $VA_c - VA_i$ comparison or across auditory areas for our $AV_c - AV_i$ comparison. The response to an ignored visual stimulus was enhanced across all visual areas when attention was directed to an auditory stimulus sharing the same side of space as the ignored visual stimulus, with complementary effects in the auditory domain. Interestingly, these effects of redirecting attention across space in a different modality also differ from the effects of redirecting attention across space in the same modality in the visual domain ($VV_c - VV_i$), where we find significant differences across early visual areas, but not in the auditory domain ($AA_c - AA_i$), where we find no significant differences across early auditory areas.

Our results confirm and extend an event-related potentials study, where both visual and auditory stimuli were presented at the same spatial location by speakers and where the visual and auditory stimuli were not presented concurrently (Eimer et al. 2004). Eimer and colleagues found that the N1, but not the P1,

component of ERP responses was enhanced in early sensory areas for ignored, task-irrelevant stimuli when the other modality attended was on the same side of space, even at a finer spatial scale within a hemifield. Here we demonstrate that such effects are present as early as V1 or A1 and that these effects do not show significant differences across the early visual or auditory areas we studied.

Although the effects of spatial attention we observed from a different modality were reduced compared with the effects of spatial attention within the same modality, it is possible that our stimuli were suboptimal. Our visual and auditory stimuli were lateralized only to the same side of space, not localized to the same spatial location, in either the frontoparallel plane or in depth (extra-personal vs. peri-personal space). The influence of location on cross-modal attention may depend on the perceptual cohesion of the visual and auditory stimuli presented. Shared spatial location is just one means of creating perceptual cohesion. Other means of manipulating the extent to which visual and auditory stimuli are perceptually linked may yield stronger effects, such as temporal synchrony/asynchrony of stimulus onset/offset or shared changes in spatial location, such as congruent visual and auditory motion. Thus it is possible that our cross-modal spatial attention effects could have been more dramatic if our auditory and visual stimuli were more strongly colocalized in space or more strongly correlated and integrated into a unified cross-modal percept.

In support of this notion, it has already been shown that the extent to which auditory information can influence visual performance depends on whether features are shared across modalities (McDonald et al. 2003; Meyer and Wuerger 2001; Shams et al. 2000; Sheth and Shimojo 2004; Watanabe and Shimojo 2001). Moreover, in multimodal areas such as VIP, many multimodal neurons have overlapping visual and auditory receptive fields, and neuronal activity is modulated by the spatial location of both visual and auditory stimuli (Schlack et al. 2005).

It remains to be seen whether the differences we observed across visual areas could be attributed to our particular choice of visual stimuli (moving gratings) or visual task (speed discrimination), both of which are optimal for area MT+. Previous work suggests that modulations in the fMRI response are task specific, with increased responses in visual areas whose neuronal selectivity best matches that of the behavioral task (Beauchamp et al. 1997; Chawla et al. 1999; Corbetta et al. 1990; Huk and Heeger 2000; O'Craven et al. 1997; Watanabe et al. 1998; but see Buračas et al. 2005). Similarly, across modalities, studies of tactile motion suggest that the visual area best suited for a task is most likely to be recruited for a corresponding task in a different modality (Ghazanfar and Schroeder 2006). Future studies will need to determine how task differences could influence the effects we observe not only in visual areas, but also in auditory areas, where, for example, spectral and temporal auditory processing might be specific not simply for a given auditory area but for the left or right hemisphere of that auditory area (e.g., Devlin et al. 2003; Zatorre and Belin 2001).

In conclusion, we measured spatial attention and cross-modal attention, and the effects of spatial attention on another modality within early auditory and visual cortex. We found that the effects of both spatial attention and cross-modal attention increased across the visual hierarchy and that cross-

modal attention effects increased from primary to secondary auditory cortex. Within the same modality, attention has been shown to enhance neuronal responses to attended locations or features and decrease responses to unattended locations or features, with the magnitude of the increase or decrease depending on cortical area. The effects we observe for cross-modal attention seem to operate in a similar manner, with effects varying across early visual and auditory areas. Thus the selective enhancement or suppression of sensory processing by cross-modal attention is likely to complement the role of attention in enhancing or suppressing processing within a single modality and, as has been found for visual attention, may depend on the nature of the features being processed and how they are shared across modalities.

The ability of auditory attention to influence visual responses, even for ignored visual stimuli, may depend on the degree to which information across modalities is perceived as originating from the same object or as sharing a common feature, or a common spatial location. In our experiment, the auditory and visual stimuli were not perceptually linked into a single object. If our stimuli had been optimized for creating a unified cross-modal percept, the cross-modal spatial attention effects we observed might have been more dramatic. Nonetheless, we found a significant enhancement when an ignored and an attended stimulus in different modalities shared the same side of space. Interestingly, our cross-modal spatial attention parameter γ did not vary across early visual or auditory areas, unlike our cross-modal or spatial attention term within the visual modality, or our cross-modal attention term within the auditory modality.

Overall, our results are consistent with the idea that lower-level visual areas provide a more veridical representation of the information presented to the retina (Treue 2003), whereas more central visual areas, such as MT+, reflect more behaviorally significant aspects of visual information. Similarly, our results are consistent with the idea that more medial auditory areas (such as A1) more faithfully represent basic auditory stimulus properties, whereas more lateral areas (such as A2) may reflect the behavioral salience of auditory information (Petkov et al. 2004). It remains to be seen whether and how the attention effects we observe in auditory cortex may vary across more detailed subdivisions of early auditory areas and how the effects of attention in visual and auditory cortex may vary based on the type of information attended or the task at hand.

ACKNOWLEDGMENTS

We thank A. Jacobs, I. Fine, A. Schlack, J. Serences, and three anonymous reviewers for helpful comments on this manuscript.

GRANTS

This work was supported by National Eye Institute Grant EY-12925 to G. M. Boynton.

REFERENCES

- Alho K, Medvedev SV, Pakhomov SV, Roudas MS, TerAniemi M, Reinikainen K, Zeffiro T, Naatanen R.** Selective tuning of the left and right auditory cortices during spatially directed attention. *Cogn Brain Res* 7: 335–341, 1999.
- Arun KS, Huang TS, Blotstein SD.** Least-squares fitting of two 3-D point sets. *IEEE Trans Pattern Anal Machine Intell* 9: 698–700, 1987.
- Baldassi S, Verghese P.** Attention to locations and features: different top-down modulation of detector weights. *J Vis* 5: 556–570, 2005.

- Beauchamp MS, Cox RW, DeYoe EA.** Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J Neurophysiol* 78: 516–520, 1997.
- Berman RA, Colby CL.** Auditory and visual attention modulate motion processing in area MT+. *Cogn Brain Res* 14: 64–74, 2002.
- Boemio A, Fromm S, Braun A, Poeppel D.** Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci* 8: 389–395, 2005.
- Boynton GM, Engel SA, Glover GH, Heeger DJ.** Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207–4221, 1996.
- Brainard DH.** The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997.
- Buračas GT, Boynton GM.** The effects of spatial attention on contrast response functions in human visual cortex. *J Neurosci* 27: 93–97, 2007.
- Buračas GT, Fine I, Boynton GM.** The relationship between task performance and fMRI response. *J Neurosci* 25: 3023–3031, 2005.
- Calvert GA, Brammer MJ, Bullmore ET, Campbell R, Iversen SD, David SA.** Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport* 10: 2619–2623, 1999.
- Calvert GA, Campbell R, Brammer MJ.** Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10: 649–657, 2000.
- Cavada C, Goldman-Rakic PS.** Posterior parietal cortex in rhesus monkey. I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *J Comp Neurol* 287: 393–421, 1989a.
- Cavada C, Goldman-Rakic PS.** Posterior parietal cortex in rhesus monkey. II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol* 287: 422–445, 1989b.
- Chambers CD, Stokes MG, Mattingley JB.** Modality-specific control of strategic spatial attention in parietal cortex. *Neuron* 44: 925–930, 2004.
- Chawla D, Rees G, Friston KJ.** The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci* 2: 671–676, 1999.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE.** Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248: 1556–1559, 1990.
- Desimone R, Gross CG.** Visual areas in the temporal cortex of the macaque. *Brain Res* 178: 363–380, 1979.
- Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J, Wandell BA.** Visual field representations and locations of visual areas V1/2/3 in human cortex. *J Vis* 3: 586–598, 2003.
- Driver J, Spence C.** Multisensory perception: beyond modularity and convergence. *Curr Biol* 10: 731–735, 2000.
- Eimer M, van Velzen J, Driver J.** ERP evidence for cross-modal audiovisual effects of endogenous attention within hemifields. *J Cogn Neurosci* 16: 272–288, 2004.
- Eimer M, van Velzen J, Forster B, Driver J.** Shifts of attention in light and in darkness: an ERP study of supramodal attentional control and crossmodal links in spatial attention. *Cogn Brain Res* 15: 308–323, 2003.
- Engel SA, Glover GH, Wandell BA.** Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 7: 181–192, 1997.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN.** fMRI of human visual cortex (Correspondence). *Nature* 369: 525, 1994.
- Falchier A, Clavagnier S, Barone P, Kennedy H.** Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci* 22: 5749–5759, 2002.
- Foxe JJ, Simpson GV, Ahlfors SP, Saron CD.** Biasing the brain's attentional set. I. Cue driven deployments of intersensory selective attention. *Exp Brain Res* 166: 370–392, 2005.
- Gandhi SP, Heeger DJ, Boynton GM.** Spatial attention affects brain activity in human primary visual cortex. *Proc Natl Acad Sci USA* 96: 3314–3319, 1999.
- Gattass R, Gross CG, Sandell JH.** Visual topography of V2 in the macaque. *J Comp Neurol* 201: 519–539, 1981.
- Gattass R, Sousa AP, Gross CG.** Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8: 1831–1845, 1988.
- Ghazanfar AA, Schroeder CE.** Is neocortex necessary? *Trends Cogn Sci* 10: 278–285, 2006.
- Hayden BY, Gallant JL.** Time course of attention reveals different mechanisms for spatial and feature-based attention in area V4. *Neuron* 47: 637–643, 2005.
- Heeger DJ, Boynton GM, Demb JB, Seidemann E, Newsome MT.** Motion opponency in visual cortex. *J Neurosci* 19: 7162–7174, 1999.

- Huk AC, Heeger DJ.** Task-related modulation of visual cortex. *J Neurophysiol* 83: 3525–3536, 2000.
- Huk AC, Heeger DJ.** Pattern-motion responses in human visual cortex. *Nat Neurosci* 5: 72–75, 2002.
- Jancke L, Mirzazade S, Shah NJ.** Attention modulates activity in the primary and secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266: 125–128, 1999.
- Jancke L, Specht K, Shah JN, Hugdahl K.** Focused attention in a simple dichotic listening task: an fMRI experiment. *Cog Brain Res* 16: 257–266, 2003.
- Johnson JA, Zatorre RJ.** Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cereb Cortex* 15: 1609–1620, 2005.
- Kastner S, DeWeerd P, Pinsk MA, Elizondo MI, Desimone R, Ungerleider LG.** Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex. *J Neurophysiol* 86: 1398–1411, 2001.
- Laurienti PJ, Burdette JH, Wallace MT, Yen, Y-F, Field AS, Stein BE.** Deactivation of sensory-specific cortex by cross-modal stimuli. *J Cogn Neurosci* 14: 420–429, 2002.
- Lewis JW, Beauchamp MS, DeYoe EA.** A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex* 10: 873–888, 2000.
- Lewis JW, Van Essen DC.** Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J Comp Neurol* 428: 112–137, 2000.
- Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S.** Attention-dependent changes of activation and connectivity in dichotic listening. *NeuroImage* 17: 643–656, 2002.
- Liu T, Stevens ST, Carrasco M.** Comparing the time course and efficacy of spatial and feature-based attention. *Vision Res* 47: 108–113, 2007.
- Macaluso E, Frith CD, Driver J.** Modulation of human visual cortex by crossmodal spatial attention. *Science* 289: 1206–1208, 2000.
- Martinez A, Anillo-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA.** Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat Neurosci* 2: 364–369, 1999.
- Martinez-Trujillo JC, Treue S.** Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14: 744–751, 2004.
- McDonald JJ, Teder-Salejarvi WA, Hillyard SA.** Involuntary orienting to sound improves visual perception. *Nature* 407: 906–908, 2000.
- Meyer GF, Wuergler SM.** Cross-modal integration of auditory and visual motion signals. *Neuroreport* 12: 2557–2560, 2001.
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ.** Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15: 963–974, 2005.
- O'Craven KM, Rosen BR, Kwong KK, Triesman A, Savoy RL.** Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* 18: 591–598, 1997.
- Pelli DG.** The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10: 437–442, 1997.
- Penhune VB, Zatorre RJ, MacDonald JD, Evans AC.** Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurements from MR scans. *Cereb Cortex* 6: 661–672, 1996.
- Petkov CI, Kang X, Alho K, Bertrand O, Yund EW, Woods DL.** Attentional modulation of human auditory cortex. *Nat Neurosci* 7: 658–663, 2004.
- Rees G, Frith C, Lavie N.** Processing of irrelevant visual motion during performance of an auditory attention task. *Neuropsychologia* 39: 937–949, 2001.
- Rockland KS, Ojima H.** Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol* 50: 19–26, 2003.
- Schlack A, Sterbing-D'Angelo SJ, Hartung K, Hoffman K-P, Bremmer F.** Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25: 4616–4625, 2005.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB.** Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995.
- Shams L, Kamitani Y, Shimojo S.** Illusions. What you see is what you hear (Brief Communication). *Nature* 408: 788, 2000.
- Shams L, Kamitani Y, Shimojo S.** Visual illusion induced by sound. *Cogn Brain Res* 14: 147–152, 2002.
- Sheth BR, Shimojo S.** Sound-aided recovery from and persistence against visual filling-in. *Vision Res* 44: 1907–1917, 2004.
- Shih SI, Sperling G.** Is there feature-based attentional selection in visual search? *J Exp Psychol Hum Percept Perform* 22: 758–779, 1996.
- Shomstein S, Yantis S.** Control of attention shifts between vision and audition in human cortex. *J Neurosci* 24: 10702–10706, 2004.
- Somers DC, Dale AM, Seiffert AE, Tootell RB.** Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc Natl Acad Sci USA* 96: 1663–1668, 1999.
- Stein BE.** Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp Brain Res* 123: 124–135, 1998.
- Strayer DL, Johnston WA.** Driven to distraction: dual task studies of simulated driving and conversing on a cellular telephone. *Psychol Sci* 12: 462–466, 2001.
- Teder-Salejarvi WA, McDonald JJ, Di Russo F, Hillyard SA.** An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cogn Brain Res* 14: 106–114, 2002.
- Teo PC, Sapiro G, Wandell BA.** Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Trans Med Imaging* 16: 852–863, 1997.
- Tervaniemi M, Hugdahl K.** Lateralization of auditory-cortex functions. *Brain Res Rev* 43: 231–246, 2003.
- Theeuwes J.** Effects of location and form cuing on the allocation of attention in the visual field. *Acta Psychologica* 72: 177–192, 1989.
- Tootell RB, Reppas JB, Dale AM, Look RB, Sereno MI, Malach R, Brady TJ, Rosen BR.** Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* 375: 139–141, 1995.
- Tootell RB, Taylor JB.** Anatomical evidence for MT and additional cortical visual areas in humans. *Cereb Cortex* 5: 39–55, 1995.
- Treue S.** Neural correlates of attention in primate visual cortex. *Trends Neurosci* 24: 295–300, 2001.
- Treue S.** Climbing the cortical ladder from sensation to perception. *Trends Cogn Sci* 7: 469–471, 2003.
- VanEssen DC, Newsome WT, Maunsell JHR.** The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Res* 24: 429–448, 1984.
- Watanabe K, Shimojo S.** When sound affects vision: effects of auditory grouping on visual motion perception. *Psychol Sci* 12: 109–116, 2001.
- Watanabe T, Harner AM, Miyauchi S, Sasaki Y, Nielsen M, Palomo D, Mukai I.** Task-dependent influences of attention on the activation of human primary visual cortex. *Proc Natl Acad Sci USA* 95: 11489–11492, 1998.
- Weissman DH, Warner LM, Woldorff MG.** The neural mechanisms for minimizing cross-modal distraction. *J Neurosci* 24: 10941–10949, 2004.
- Westbury CF, Zatorre RJ, Evans AC.** Quantifying variability in the planum temporale: a probability map. *Cereb Cortex* 9: 392–405, 1999.
- Williford T, Maunsell JH.** Effects of spatial attention on contrast response functions in macaque area V4. *J Neurophysiol* 96: 40–54, 2006.
- Woldorff MG, Gallen CC, Hampson SA, Hillyard SA.** Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci USA* 90: 8722–8726, 1993.
- Zatorre RJ, Belin P.** Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11: 946–953, 2001.
- Zatorre RJ, Mondor TA, Evans AC.** Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage* 10: 544–554, 1999.